AD_____

Award Number: MIPR 3GD3E93093

TITLE: Phase II-SOF Knowledge Coupler-Based Phase I XML Schema

PRINCIPAL INVESTIGATOR: Colonel Warren L. Whitlock

CONTRACTING ORGANIZATION: Southeast Regional Medical Command
Fort Gordon, GA 30905-5650

REPORT DATE: March 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-03-2005	Final	17 Mar 03 - 30 Sep 03

4. TITLE AND SUBTITLE
Phase II-SOF Knowledge Coupler-Based Phase I XML Schema

5a. CONTRACT NUMBER

5b. GRANT NUMBER
MIPR 3GD3E93093

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Colonel Warren L. Whitlock

E-Mail: warren.whitlock@se.amedd.army.mil

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Southeast Regional Medical Command
Fort Gordon, GA 30905-5650

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

See attached.

15. SUBJECT TERMS

No subject terms provided.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	9	

Table of Contents

Phase II-SOF Knowledge Coupler-Based Phase I XML Schema
Final Report
6 March 2005

ABSTRACT:

The 2002 digital version of the Special Operations Forces Medical Handbook (SOFMH) includes a comprehensive, searchable database of injuries and illnesses. While it is a complete digital reference source, its utility would be greatly enhanced if a healthcare provider could enter a patient's signs and symptoms into the SOFMH search template and access a list of diagnostic choices in an XML-tagged database.

An analysis of the search function indicates that the native search capability of the SOFMH does not inherently contain the requirements to sustain a diagnostic tool. Current search technologies can locate text or indexes, ranked by the frequency a term appears in a document, but not the term's relevance to a set of symptoms. Current search technologies operate by diagnosis category, key words, indices, and content text. The program ranks matches by frequency, index, and content. A medical knowledge coupler requires more sophisticated associations to link a diagnosis to the symptom. XML tagging was selected as the method to identify and assign significance to portions of text information. Initial tagging of the SOFMH did not enable the level of detail required for the diagnostic process. The next step towards a reliable diagnostic tool is establishing the relationships between the symptoms and the diagnosis. A diagnostic tool cannot automatically make these associations; it must be provided the information. Keywords will be used prescribe a certain intuitiveness in the application. A keyword can be significant in many different diagnoses, but will have different weighted values depending on its association with other keywords in a symptom set. Further, the overall weight or ranking of a particular diagnosis in relation to other diagnoses may change or be thrown out completely due to other factors. Collaboration with the Stanford University School of Medicine Department of Medical Informatics on the Stanford XML Tagging Tool generated a web-based architecture, but it is not operational for remote or stand-alone computer users. Improving diagnostic accuracy and inserting discrimination capability to allow for sophisticated tagging of information concerning adverse drug effects, focused history questions, or assess to all specific requirements for long-term disposition needed for profiling recommendations and follow-up.

Developing the relationships will require creation of a database tagging system to establish and test the relationships. This will provide specific diagnosis, linked to care plans, based on operational circumstances that can be inserted into the medical record. Database queries will be used to validate relationships, forming the basis for a client-side knowledge-coupler. Database queries will be used to test the validity of the relationships using a web interface, forming the basis for a web-enabled diagnostic tool. A stand-alone application will be developed for use in environments where client-side applications are required, e.g., (forward locations, Personal Digital Assistants, or unstable web connectivity).

FINAL RESULTS:

The SOFMKC is a concept that has the potential to revolutionize military healthcare, especially in austere environments or remote locations where enhanced medical support is lacking. However, for a tool to be constructed using this technology and for it to be useful, it must be both portable and simple to operate. Technology is available to make a portable device, medical treatment policies and procedures are available, a robust user interface is required to link the two. Simplification of operation requires a sophisticated methodology backbone to be flexible and accurate.

The Special Operations Forces Medical Handbook (SOFMH) content was used as a basis for the content. An initial text parser was built that placed text into a structured database using text source files. This database provided the original information to display to SOF medics or editors using the application. Additional database development is required to further isolate and refine the information into a delivery format that would target the specificity required to assist in determining treatment options.

The Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT) Database developer license was purchased from the College of American Pathologists. This license was necessary to obtain nomenclature relationship files.

The use of the SNOMED-CT database ensured that the information is highly organized and available. The central issue is determining the subset organization applicable to a combat medic, which requires a high level of effort to realize. This level of effort failed to materialize due to many factors (as discussed in the Problems Encountered section).

Several database applications were created to allow manipulation and display of medical knowledge over the complete SOFMH content.

- An image database was created to manage the 700+ images included in the CD-ROM version of the SOFMH. This database includes assignments of the College of American Pathologists' Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT) codes to particular images. Images were described by as many different codes as possible to enable the medic to access the image. Multiples SNOMED-CT codes were assigned to each image according to the anatomic location, organ system, medical specialty area, operational environment, trauma, and procedures. All animals and microscopic organisms were assigned the SNOMED-CT codes for its class, order, family, genus, and species. These images were compiled into an online catalog as well as appearing on their original SOFMH CD-ROM topic page (electronic edition). This database demonstrated "proof-of-concept" to establish the relationships between SNOMED-CT codes and an image could be linked in numerous diagnoses or conditions based on other factors in the patient's presentation of symptoms.

- Comprehensive Analysis of the SNOMED-CT database architecture. A SNOMED-CT database browser was created to traverse the extensive relationships for symptom analysis and query formulation for the knowledge coupler engine.

- A web-enabled database editing application was completed to allow online updating and virtual collaboration of geographically separated authors and subject matter experts to meet the author and editorial review board requirements.

- Created a relationships database for use in the Special Operations Forces Medical Knowledge Coupler (SOFMKC). This database contains the original content, SQL scripts, and stored procedures to manage the interactivity of the application as well as the source data for the knowledge coupler display. This database is necessary to provide the linkages between the navigation hierarchy of the SNOMED-CT concepts and the signs and symptoms indicated in a patient interview. This product included several versions of an anatomical patient interview created by a medical illustrator for this purpose.

- A separate web tool was written to display the relationships between distinct SNOMED-CT concepts. This tool was necessary to determine the scope of each concept ID and its relationship to other concept IDs, contrived through the linkage of other concept IDs that are used to determine the type of relationship of the two original concept IDs. Examples of these relationships might be parent, child or sibling, etc.

These applications were necessary to allow authors to study the clinical architecture and build associations necessary to the formulation of the relationships integral to symptom versus diagnosis suggestion. Author and subject matter expert input was required to maintain the suitability of the suggested treatment to the capabilities of the SOF medic. The formation of the editorial board for the second edition of the SOFMH was to be the basis of the subject matter expert's content to be applied to the SOFMKC.

DELIVERABLE:

A stand-alone tool was not created, however, several applications integral to the effort (detailed above) were completed including;

- An College of American Pathologists' Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT) image database.

- SNOMED-CT database browser.

- Web-enabled relationships database to link symptoms with specificity.

- Online editorial tool

- Anatomical symptom tool

PROBLEMS ENCOUNTERED:

When the original SOFMH digital version was created in 2002, it was the goal of the Special Operations Command (SOCOM) Surgeon to do yearly updates of the text and the media. This would require a high level of effort on the part of SOCOM physicians, physician assistants, and SOF medics as authors and editors. Because of lack of funding, as well as the increased operations tempo created by the wars in Afghanistan and Iraq, plans by the SOCOM Surgeon to update the SOFMH had to be put on hold for an indefinite period of time. Since it was essential for the symptomatology to diagnosis data linkages for SOFMKC to be done by the subject matters experts (SME) as they updated/rewrote their individual sections, a critical piece of the SOFMKC development was missing because the SMEs were unable to participate. The collaboration necessary to assemble subject matter experts and maintain focus on this project in the face of extremely busy case loads and difficulties in travel and scheduling conflicts contributed to this problem.

The clinical symptoms and robust relationship codes and modifiers within SNOMED-CT are not complete. Relationships do not exist for many of the symptoms and signs that need to be available for the medics to adequately meet the demands of emergency medicine to the extent required in a first responder in primitive environments where immediate transfer of care is not available. Subsequent SOFMH updates must complete these relationships, to include primitive treatment options, geographic location indicators to typify disease probability, and image relationships (where available) to reduce the time factor in diagnoses by more rapidly eliminating unlikely possibilities.

Both of the above issues led to the need for an online editorial application. This online application requires high security to authenticate the user, thereby maintaining the audit trail accounting of information and credible source content, and the ability to conference on particular topics. It also requires the creation of an editorial board role to both solicit input and limit updating of information. Where differing opinions create deadlock, the editor board will need to be able to halt further editing or prevent out-of-cycle changes to the content.

Current technologies cannot adequately store the amount of information required to be available in a hand-held device using a stand-alone application. The SNOMED-CT nomenclature alone contains over 300,000 unique classifications. In addition to the processing required to locate the proper SNOMED-CT code structure, the database engine necessary to locate the SNOMED-CT code that provides the base relationship will require processing resources. Besides these resources other applications, such as the image database, will require processing and memory allocation.

SNOMED-CT codes are associated with each other through a complicated algorithm and contained in a subset members table. A subset editor has been recently been released, but a subset still needs to be built to the specifications of the combat medic level of care. The creation of this subset requires customizing efforts that are tedious, complicated, and require a much more in-depth knowledge of the architecture than is currently available within our current resource and task level limits.

During the creation of the applications, a newer version of the SNOMED-CT database was received. No information was supplied with it regarding merging versions. We

developed a manual system to import the revised database information. Several database tables did not match the technical descriptions causing errors that required troubleshooting to ensure a successful implementation of the database content.

Research determined that an XML Database would not provide any additional benefit over a traditional SQL database in this scenario, so none was developed. However, an XML schema was developed and tested during the concept phase. Using a strict XML construction provided the desired portability, but was determined to be unsuitable due to the lengthy search time required to parse the XML tree structure and deliver a response, therefore further development was halted at this stage in the project.

Since an initial prototype portable SOFMKC was not completed, no evaluation of the product by medical personnel was possible.

AMEDD-WIDE ADOPTION POTENTIAL:

A working SOFMKC does not exist at this time, but the Online Editor tool developed as a component of the application could be used to update any collaborative project that requires geographically separate individuals with web connectivity to interact.

NEXT STEPS:

The data parser must be updated to further refine the transformed raw data into the SOFMKC API format and revise the SOFMKC. This will provide for rapid revision of the SOFMH in addition to presenting the knowledge coupler data in a standardized, familiar format for ease of use.

Clinical queries need to be created and validated for solution set to data to be indicative of symptoms that are most likely to be observed by the SOF medic, rather than including complex, highly technical treatments that do not involve initial contact, life support activities, but would be accomplished at rear echelon, larger medical facilities. This will serve to reduce the solution set to a more manageable size and prevent human factor issues such as raised frustration levels and maintain the viability and specificity of the tool.

Functionality of the image catalog database must be verified. Image catalog queries must be incorporated into the main database engine. An additional feature could be adding the images as another clue to the diagnosis by displaying possible image thumbnails to compare with the presented symptoms, linked via the SNOMED-CT code to the result set.

We must continue to incorporate the SNOMED-CT nomenclature to classify the information and provide "handles" to the data. Additionally, we would need to build pre-query searches with text and technical handles and have the query search the handles rather than performing either a full text (pre-limited by searching only specific database text columns) or SNOMED-CT database search. The pre-query data handles can determine which results to display depending on an algorithm that combines the number and type handles to determine and rank the result set.

Metrics need to be developed for evaluating students' performance while using web assistance applications in the Combat Trauma Patient Simulator against a control group who will not have access to web assistance applications. The efficiency of the tool can also be tested and further refinements to the interface accomplished.

We would need to create metric scenarios for test evaluation of cases by medical personnel. This would be done by using the web application to conduct additional user interface operability study to locate issues raised by the use of the application. This will further enhance the design of the user interface and pre-query efficiencies (a student will use an application differently than an experienced technician). Additionally, search logic will be tested, and additional clues built as needed to assist in the formulation of proper symptom information to determine accuracy of the presentation.

Once a successful web application has been created and the median limits of the database solution set established, a basic template application for a Personal Digital Assistant (PDA) will be developed and tested. The database set for the PDA will include a streamlined database; overly specific, advanced treatment options will be removed.

Integration of some of the SOFMKC components or the SOFMKC end-product into BMIS-T would be a final goal.

CONCLUSIONS:

Completion of the SOFMKC would provide an extremely useful tool for first responders at all levels of treatment. Medical knowledge continues to advance and updating information stores is inevitable. Coupling this project with future updates of the SOFMH as conceived would be advantageous for both efforts and pay dividends in many aspects of operational medicine.